

Response Factors of VOC Analyzers at
a Meter Reading of 10,000 ppmv for
Selected Organic Compounds

Radian Corp.
Austin, TX

Prepared for

Industrial Environmental Research Lab.
Research Triangle Park, NC

Mar 81



EPA-600/2-81-051

August 1981

PB81-234817

RESPONSE FACTORS OF VOC ANALYZERS
AT A METER READING OF 10,000 PPMV
FOR SELECTED ORGANIC COMPOUNDS

by

D.A. DuBose and G.E. Harris

Radian Corporation
P.O. Box 9948
Austin, Texas 78766

Contract No. 68-02-3171-28

EPA Project Officer: Dr. B.A. Tichenor

Chemical Processes Branch
EPA/IERL
Research Triangle Park, NC 27711

Prepared for

Office of Air Quality Planning and Standards

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA 22161

TECHNICAL REPORT DATA (Please read Instructions on the reverse before completing)		
1. REPORT NO. EPA-600/2-81-051	2.	3. RECIPIENT'S ACCESSION NO. PBM 234817
4. TITLE AND SUBTITLE Response Factors of VOC Analyzers at a Meter Reading of 10,000 ppmv for Selected Organic Compounds	5. REPORT DATE March 1981	
7. AUTHOR(S) D. A. DuBose and G. E. Harris	6. PERFORMING ORGANIZATION CODE	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Radian Corporation P.O. Box 9948 Austin, Texas 78766	8. PERFORMING ORGANIZATION REPORT NO.	
12. SPONSORING AGENCY NAME AND ADDRESS EPA, Office of Research and Development Industrial Environmental Research Laboratory Research Triangle Park, NC 27711	10. PROGRAM ELEMENT NO. 1AB604	
	11. CONTRACT/GRANT NO. 68-02-3171, Task 28	
15. SUPPLEMENTARY NOTES IERL-RTP project officer is Bruce A. Tichenor, Mail Drop 62, 919/541-2547. EPA-600/2-81-002 is a related report.	13. TYPE OF REPORT AND PERIOD COVERED Task Final; 12/80-2/81	
	14. SPONSORING AGENCY CODE EPA/600/13	
16. ABSTRACT The report summarizes results of a reinterpretation of data generated in a laboratory study of the sensitivity of two types of portable hydrocarbon detectors to a variety of organic chemicals. (A previous report, EPA-600/2-81-002, describes and gives original results of the laboratory study.) Detector sensitivity is quantified by a response factor for each chemical where the response factor equals the actual concentration of the chemical divided by the observed concentration from the detector. The previous report estimated response factors at 10,000 ppmv actual concentration of the chemical. This report presents response factors estimated for a 10,000 ppmv detector reading. The instruments were calibrated to 7993 ppmv methane gas.		
17. KEY WORDS AND DOCUMENT ANALYSIS		
a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Pollution Hydrocarbons Measuring Instruments Sensitivity Organic Compounds	Pollution Control Stationary Sources Hydrocarbon Detectors Response Factors	13B 07C 14B 14G
19. DISTRIBUTION STATEMENT Release to Public	19. SECURITY CLASS (This Report) Unclassified	21. NO. OF PAGES
	20. SECURITY CLASS (This page) Unclassified	22. PRICE

TABLE OF CONTENTS

<u>Section</u>	<u>Page</u>
List of Tables	ii
1. INTRODUCTION	1
2. RESULTS	2
3. DATA ANALYSIS.	12
3.1 Classical Estimation of Response Factors	12
3.2 Inverse Estimation of Response Factors	14
3.3 Discussion	15
4. LITERATURE REVIEW.	17
References	24

LIST OF TABLES

<u>Number</u>		<u>Page</u>
2-1	Response Factors with 95% Confidence Intervals Estimated at 10,000 ppmv Response	3
2-2	A List of Tested Chemicals which appear to be unable to Achieve An Instrument Response of 10,000 ppmv at Any Feasible Concentration	11

SECTION 1

INTRODUCTION

This technical note summarizes the results of a reinterpretation of the data generated in a laboratory study of the sensitivity of two types of portable hydrocarbon detectors to a variety of organic chemicals. This work is funded by the EPA as part of Contract Number 68-02-3171, Task 28. A previous report (Brown, et al., 1980) presents the description and original results of the laboratory study.

The detector sensitivity is quantified by a "response factor" for each chemical where

$$\text{Response Factor} = \frac{\text{Actual Concentration of Chemical}}{\text{Observed Concentration from Detector}}$$

The previous report (Brown, et al., 1980) estimated response factors of 10,000 ppmv actual concentration of the chemical. This report presents response factors estimated for a 10,000 ppmv detector reading.

The instruments were calibrated to 7993 ppmv methane gas.

SECTION 2

RESULTS

This Technical Note presents response factors for two types of portable hydrocarbon analyzers, the "OVA-108" and the "TLV sniffer", for 168 different chemical compounds. The response factor is defined as the ratio of the actual concentration to the observed concentration (or instrument response).

The response factor varies with concentration. In a previous report (Brown, et al., 1980) the response factors were presented at 10,000 ppmv actual concentration of the chemical.

The response factors are computed at 10,000 ppmv concentration observed by the detector in this Technical Note. Table 2-1 presents response factors for chemicals on both types of instruments along with the 95 percent confidence intervals. The Background Information Document on fugitive emissions from the Synthetic Organic Chemicals Manufacturing Industry (SOCMI) (EPA, 1980) recommends an action level of 10,000 ppmv as read directly on a portable hydrocarbon detector.

Two statistical methods were used here to compute estimates and associated confidence intervals: the classical regression method and the inverse regression method (See Section 3). Both methods of estimation essentially involve the fitting of a line to the actual and instrument observed concentrations. The interpolated or extrapolated actual concentration at 10,000 ppmv observed concentration is calculated based on the fitted line. The response factor is computed as the ratio of this estimated actual concentration to the specified 10,000 ppmv "observed" concentration.

TABLE 2-1

RESPONSE FACTORS
WITH 95 % CONFIDENCE INTERVALS
ESTIMATED AT 10,000 PPMV RESPONSE

OCPDB* ID NO	COMPOUND NAME	VOLATILITY CLASS**	OVA		TLV	
			RESPONSE FACTOR	CONFIDENCE INTERVALS	RESPONSE FACTOR	CONFIDENCE INTERVALS
70	ACETIC ACID	LL	1.64	(1.11, 2.85)	15.90	(7.05, 46.20)
80	ACETIC ANHYDRIDE	LL	1.39	(1.09, 1.88)	5.88 I	(2.71, 12.80)
90	ACETONE	LL	0.80	(0.57, 1.20)	1.22	(0.81, 2.00)
100	ACETONE CYANOHYDRIN	HL	3.51	(0.89, >100.00)	21.00 N	(1.09, >100.00)
110	ACETONITRILE	LL	0.95	(0.85, 1.06)	1.18	(0.94, 1.52)
120	ACETOPHENONE	HL	18.70	(5.52, >100.00)	B	(1.65, 5.32)
125	ACETYL CHLORIDE	LL	2.04	(1.72, 2.48)	2.72	(1.65, 5.32)
130	ACETYLENE	G	0.39	(0.38, 0.43)	B	(0.44, 27.90)
180	ACRYLIC ACID	LL	4.59	(3.38, 6.57)	B	(0.44, 27.90)
170	ACRYLONITRILE	LL	0.97	(0.80, 1.20)	3.49 I	(0.44, 27.90)
	ALLENE	G	0.64	(0.60, 0.69)	15.00 X	(9.68, 26.50)
200	ALLYL ALCOHOL	LL	0.98	(0.78, 1.27)		
250	AMYL ALCOHOL, N-	HL	0.75	(0.57, 1.04)	2.14	(0.45, >100.00)
2855	ANYLENE	LL	0.44	(0.34, 0.61)	1.03	(0.59, 2.59)
320	ANISOLE	LL	0.92	(0.65, 1.48)	3.91	(0.52, >100.00)
360	BENZALDEHYDE	HL	2.48	(1.38, 5.62)	B	(0.96, 1.20)
	BENZENE	LL	0.29	(0.28, 0.31)	1.07	(0.96, 1.20)
450	BENZONITRILE	HL	2.99	(1.18, 15.30)	B	
490	BENZOYL CHLORIDE	HL	22.10 D	(3.43, >100.00)	B	
530	BENZYL CHLORIDE	HL	15.30 D	(3.98, >100.00)	B	
570	BROMOBENZENE	LL	0.40	(0.34, 0.48)	1.19	(0.27, >100.00)
590	BUTADIENE, 1,3-	G	0.57	(0.54, 0.60)	10.90	(8.11, 15.40)
	BUTANE, N-	G	0.50	(0.48, 0.55)	0.83	(0.58, 0.70)
640	BUTANOL, N	LL	1.44 I	(0.89, 2.34)	4.11 I	(2.18, 7.83)
650	BUTANOL, SEC-	LL	0.78	(0.70, 0.83)	1.25	(0.99, 1.66)
660	BUTANOL, TERT	S	0.53	(0.38, 0.81)	2.17	(1.34, 4.43)

* ORGANIC CHEMICAL PRODUCERS DATA BASE

** G=GAS; LL=LIGHT LIQUID; HL=HEAVY LIQUID.

DEFINITION OF EXPLANATORY DATA CODES:

- : INVERSE ESTIMATION METHOD X NO DATA AVAILABLE
 D POSSIBLE OUTLIERS IN DATA B 10,000 PPMV RESPONSE UNACHIEVABLE
 N NARROW RANGE OF DATA P SUSPECT POINTS ELIMINATED

TABLE 2-1 (Continued)

RESPONSE FACTORS
WITH 95 % CONFIDENCE INTERVALS
ESTIMATED AT 10,000 PPBV RESPONSE

OCPDB+ ID NO	COMPOUND NAME	OVA			TLV		
		VOLATILITY CLASS**	RESPONSE FACTOR	CONFIDENCE INTERVALS	RESPONSE FACTOR	CONFIDENCE INTERVALS	
592	BUTENE, 1-	G	0.58	(0.51, 0.62)	5.84	(4.20, 8.69)	
600	BUTYL ACETATE,	LL	0.66	(0.54, 0.83)	1.38	(1.15, 1.70)	
630	BUTYL ACRYLATE, N-	LL	0.70	(0.63, 0.78)	2.57 I	(1.17, 5.68)	
	BUTYL ETHER, N	LL	2.80	(0.81, 95.60)	3.58 I	(1.82, 7.04)	
	BUTYL ETHER, SEC	LL	0.35	(0.21, 0.95)	1.15	(0.75, 2.17)	
670	BUTYLAMINE, N-	LL	0.69	(0.53, 0.98)	2.02	(1.14, 4.97)	
680	BUTYLAMINE, SEC-	LL	0.70	(0.58, 0.87)	1.56	(0.77, 5.24)	
690	BUTYLAMINE, TERT-	LL	0.63	(0.58, 0.70)	1.85	(1.42, 2.81)	
	BUTYLBENZENE, TERT-	HL	1.32	(0.89, 2.20)	2.30	(0.98, 12.80)	
750	BUTYRALDEHYDE N-	LL	1.29	(1.07, 1.61)	10.70 I	(8.53, 17.60)	
760	BUTYRIC ACID	HL	0.80	(0.38, 3.14)	1.47 I	(0.82, 3.48)	
780	BUTYRONITRILE	LL	0.52	(0.40, 0.74)	3.92	(1.87, 12.80)	
790	CARBON DISULFIDE	LL	B		5.07	(3.08, 9.78)	
830	CHLOROACETALDEHYDE	LL	8.10	(5.73, 18.20)	0.88	(0.77, 1.00)	
890	CHLOROBENZENE	LL	0.38	(0.32, 0.47)	3.90 P	(1.58, 14.10)	
1740	CHLOROETHANE	G	5.38 I	(1.87, 28.40)	18.30 I	(8.50, 51.50)	
930	CHLOROETHANE	LL	9.28	(5.19, 20.00)	0.87	(0.68, 1.18)	
960	CHLOROFORM	HL	4.58	(1.72, 27.20)	1.24	(1.08, 1.42)	
	CHLOROPHENOL, 0-	LL	0.67	(0.61, 0.73)	0.91	(0.40, 7.47)	
210	CHLOROPROPENE, 1-	LL	0.80	(0.72, 0.90)	1.08	(0.33, >100.00)	
970	CHLOROPROPENE, 3-	LL	0.48	(0.45, 0.51)	1.08	(0.77, 1.77)	
980	CHLOROTOLUENE, M-	LL	0.48	(0.42, 0.55)	4.38 I	(0.40, 47.40)	
990	CHLOROTOLUENE, O-	LL	0.58	(0.52, 0.61)	B		
	CHLOROTOLUENE, P-	S	0.98	(0.70, 1.45)			
1010	CRESOL, 0-	LL	1.25	(0.82, 2.24)			
1040	CROTONALDEHYDE	LL					

* ORGANIC CHEMICAL PRODUCERS DATA, BASE
** G=GAS; LL=LIGHT LIQUID; HL=HEAVY LIQUID.

DEFINITION OF EXPLANATORY DATA CODES:

I INVERSE ESTIMATION METHOD X NO DATA AVAILABLE
D POSSIBLE OUTLIERS IN DATA B 10,000 PPBV RESPONSE UNACHIEVABLE
N NARROW RANGE OF DATA P SUSPECT POINTS ELIMINATED

TABLE 2-1 (Continued)

RESPONSE FACTORS
WITH 95 % CONFIDENCE INTERVALS
ESTIMATED AT 10,000 PPWV RESPONSE

OCPDB* ID NO	COMPOUND NAME	VOLATILITY CLASS**	OVA			TLV		
			RESPONSE FACTOR	CONFIDENCE INTERVALS	RESPONSE FACTOR	CONFIDENCE INTERVALS	RESPONSE FACTOR	CONFIDENCE INTERVALS
1080	CUMENE	LL	1.87	(1.10, 3.71)	B	(0.62, 0.80)		
1120	CYCLOHEXANE	LL	0.47	(0.39, 0.58)				
1130	CYCLOHEXANOL	HL	0.85	(0.65, 1.20)	B			
1140	CYCLOHEXANONE	LL	1.50	(0.87, 2.76)				
1150	CYCLOHEXENE	LL	0.49	(0.42, 0.57)				
1160	CYCLOHEXYLAMINE	LL	0.57	(0.42, 0.86)				
	DECANE	HL	0.09 N	(0.05, >100.00)	I			
1190	DIACETONE ALCOHOL	HL	1.45	(0.98, 2.48)				
	DIACETYL	LL	1.54	(1.25, 1.92)				
1270	DICHLORO-1-PROPENE, 2,3-	LL	0.75	(0.58, 1.09)				
1215	DICHLOROBENZENE, M-	HL	0.64	(0.55, 0.77)				
1216	DICHLOROBENZENE, O-	HL	0.68	(0.47, 1.11)				
	DICHLOROETHANE, 1,1-	LL	0.78	(0.62, 1.02)				
1244	DICHLOROETHANE, 1,2-	LL	0.95	(0.77, 1.22)				
1235	DICHLOROETHYLENE, CIS 1,2-	LL	1.27	(1.05, 1.56)				
1238	DICHLOROETHYLENE, TRANS 1,2	LL	1.11	(0.98, 1.27)				
2620	DICHLOROMETHANE	LL	2.81	(2.13, 3.87)				
3110	DICHLOROPROPANE, 1,2-	LL	1.03	(0.82, 1.33)				
1440	DIISOBUTYLENE	LL	0.35	(0.29, 0.44)				
1870	DIMETHOXY ETHANE, 1,2-	LL	1.22	(0.84, 3.61)				
1490	DIMETHYLFORMAMIDE, N,N-	LL	4.19	(2.90, 8.58)				
1495	DIMETHYLHYDRAZINE 1,1-	LL	1.03	(0.77, 1.45)				
1520	DIMETHYLSULFOXIDE	HL	0.07 I	(0.05, 0.11)				
1480	DIOXANE	LL	1.48	(1.04, 2.33)				
1650	EPICHLOROHYDRIN	LL	1.69	(1.58, 1.84)				

* ORGANIC CHEMICAL PRODUCERS DATA BASE

** G=GAS; LL=LIGHT LIQUID; HL=HEAVY LIQUID.

DEFINITION OF EXPLANATORY DATA CODES:

I INVERSE ESTIMATION METHOD X NO DATA AVAILABLE
 D POSSIBLE OUTLIERS IN DATA B 10,000 PPWV RESPONSE UNACHIEVABLE
 N NARROW RANGE OF DATA P SUSPECT POINTS ELIMINATED

TABLE 2-1 (Continued)

RESPONSE FACTORS
WITH 95 % CONFIDENCE INTERVALS
ESTIMATED AT 10,000 PPMV RESPONSE

OCPCB* ID NO	COMPOUND NAME	VOLATILITY CLASS**	OVA		TLV	
			RESPONSE FACTOR	CONFIDENCE INTERVALS	RESPONSE FACTOR	CONFIDENCE INTERVALS
1650	ETHANE	G	0.65	(0.44, 1.58)	0.69 I	(0.21, 2.30)
1910	ETHANOL	LL	1.78	(1.59, 2.01)	X	
1910	ETHOXY ETHANOL, 2-	LL	1.55	(1.28, 1.98)	1.82	(0.86, 5.12)
1670	ETHYL ACETATE	LL	0.88	(0.77, 0.95)	1.43	(1.07, 2.00)
1880	ETHYL ACETOACETATE	HL	3.82	(1.89, 10.70)	5.80 X	(1.93, 38.80)
1690	ETHYL ACRYLATE	LL	0.77	(0.63, 0.97)		
1750	ETHYL CHLOROACETATE	LL	1.99	(1.70, 2.38)	1.59	(0.40, >100.00)
1990	ETHYL ETHER	LL	0.97	(0.77, 1.30)	1.14	(0.94, 1.42)
1710	ETHYLBENZENE	LL	0.73	(0.52, 1.11)	4.74 D	(1.38, 81.30)
1770	ETHYLENE	LL	0.71	(0.63, 0.82)	1.58	(1.26, 2.06)
1980	ETHYLENE OXIDE	G	2.48	(1.95, 3.29)	2.40	(0.88, >100.00)
1800	ETHYLENEDIAMINE	LL	1.73	(1.29, 2.48)	3.28	(0.78, >100.00)
2060	FORMIC ACID	LL	14.20	(10.60, 19.80)	B	
2105	GLYCIDOL	LL	8.88	(3.33, 19.70)	5.88	(2.08, 34.70)
	HEPTANE	LL	0.41 I	(0.28, 0.60)	0.73	(0.33, 8.10)
	HEXANE, N-	LL	0.41	(0.38, 0.45)	0.88	(0.63, 0.76)
	HEXENE, 1-	LL	0.49	(0.39, 0.68)	4.69 D	(0.85, >100.00)
	HYDROXYACETONE	LL	8.90	(4.45, 12.10)	15.20	(8.11, 88.40)
2200	ISOBUTANE	G	0.41	(0.29, 1.04)	0.55	(0.41, 0.81)
	ISOBUTYLENE	G	3.13	(0.90, 38.50)	B	
2350	ISOPRENE	LL	0.59	(0.48, 0.80)	X	
2360	ISOPROPANOL	LL	0.91	(0.72, 1.20)	1.39	(0.94, 2.31)
2370	ISOPROPYL ACETATE	LL	0.71	(0.62, 0.83)	1.31	(1.04, 1.72)
2390	ISOPROPYL CHLORIDE	LL	0.68	(0.60, 0.77)	0.98	(0.82, 1.22)
	ISOVALERALDEHYDE	LL	0.64	(0.57, 0.74)	2.19 D	(1.14, 8.85)

* ORGANIC CHEMICAL PRODUCERS DATA BASE

** G-GAS; LL-LIGHT LIQUID; HL-HEAVY LIQUID.

DEFINITION OF EXPLANATORY DATA CODES:

I INVERSE ESTIMATION METHOD X NO DATA AVAILABLE
 D POSSIBLE OUTLIERS IN DATA B 10,000 PPMV RESPONSE UNACHIEVABLE
 N NARROW RANGE OF DATA P SUSPECT POINTS ELIMINATED

TABLE 2-1 (Continued)

RESPONSE FACTORS
WITH 95 % CONFIDENCE INTERVALS
ESTIMATED AT 10,000 PPWV RESPONSE

OCFDB* ID NO	COMPOUND NAME	VOLATILITY CLASS**	OVA			TLV		
			RESPONSE FACTOR	CONFIDENCE INTERVALS	RESPONSE FACTOR	CONFIDENCE INTERVALS	RESPONSE FACTOR	CONFIDENCE INTERVALS
2450	MESITYL OXIDE	LL	1.09	(0.94, 1.29)	3.14	(1.43, 12.00)		
	METHACROLEIN	LL	1.20	(0.90, 1.71)	3.49 D	(1.51, 19.80)		
2480	METHACRYLIC ACID	HL	0.82	(0.31, 14.70)	1.08 I	(0.24, 4.58)		
2500	METHANOL	LL	4.39 P	(3.81, 5.80)	2.01	(1.66, 2.48)		
1930	METHOXY-ETHANOL, 2-	LL	2.25	(1.62, 3.34)	3.13	(1.13, 27.40)		
2510	METHYL ACETATE	LL	1.74	(1.48, 2.13)	1.85	(1.44, 2.49)		
	METHYL ACETYLENE	G	0.81	(0.58, 0.84)	6.79	(4.86, 10.40)		
2560	METHYL CHLORIDE	G	1.44	(1.22, 1.76)	1.84	(0.73, >100.00)		
2640	METHYL ETHYL KETONE	LL	0.64	(0.51, 0.84)	1.12	(0.93, 1.38)		
2645	METHYL FORMATE	LL	3.11	(2.42, 4.14)	1.94	(1.72, 2.21)		
2665	METHYL METHACRYLATE	LL	0.99	(0.90, 1.10)	2.42	(1.39, 5.38)		
2650	METHYL-2-PENTANOL, 4-	LL	1.68	(1.27, 2.32)	2.00	(1.40, 3.18)		
2660	METHYL-2-PENTANONE, 4-	LL	0.58	(0.46, 0.69)	1.63	(1.22, 2.36)		
2550	METHYL-3-BUTYN-2-OL, 2	LL	0.59	(0.44, 0.86)	X			
	METHYLAL	LL	1.37	(1.06, 1.83)	1.46	(1.24, 1.74)		
2540	METHYLANILINE, N-	HL	4.64	(3.91, 5.57)	9.48 I	(2.55, 35.20)		
2570	METHYLCYCLOHEXANE	LL	0.48	(0.28, 1.39)	0.84	(0.68, 1.09)		
	METHYLCYCLOHEXENE, 1-	LL	0.44	(0.36, 0.54)	2.79	(1.78, 5.12)		
2670	METHYLPENTYNOL	LL	1.17	(0.71, 2.48)	3.42	(1.63, 8.54)		
2690	METHYLSTYRENE, A-	LL	13.90	(8.50, 21.50)	B			
2700	MORPHOLINE	LL	0.92	(0.87, 1.40)	2.59 I	(0.84, 10.50)		
2770	NITROBENZENE	HL	1.40	(1.20, 1.65)	0.01 I	(0.00, 82.80)		
2780	NITROETHANE	LL	3.52	(3.03, 4.15)	3.45	(1.58, 13.00)		
2791	NITROMETHANE	LL	1.05	(0.80, 1.48)	7.60	(1.91, >100.00)		
2795	NITROPROPANE	LL			2.02	(1.17, 4.47)		

* ORGANIC CHEMICAL PRODUCERS DATA BASE

** G-GAS; LL-LIGHT LIQUID; HL-HEAVY LIQUID.

DEFINITION OF EXPLANATORY DATA CODES:

I INVERSE ESTIMATION METHOD X NO DATA AVAILABLE
 D POSSIBLE OUTLIERS IN DATA B 10,000 PPWV RESPONSE UNACHIEVABLE
 N NARROW RANGE OF DATA P SUSPECT POINTS ELIMINATED

TABLE 2-1 (Continued)

OCPOB* ID NO	COMPOUND NAME	VOLATILITY CLASS**	RESPONSE FACTORS WITH 95 % CONFIDENCE INTERVALS ESTIMATED AT 10,000 PPBV RESPONSE				TLV			
			OVA		TLV		RESPONSE FACTOR	CONFIDENCE INTERVALS	RESPONSE FACTOR	CONFIDENCE INTERVALS
			RESPONSE FACTOR	CONFIDENCE INTERVALS	RESPONSE FACTOR	CONFIDENCE INTERVALS				
	NONANE-N	LL	1.54	(0.94, 2.98)	11.10	(3.13, >100.00)				
	OCTANE	LL	1.03	(0.89, 1.21)	2.11	(1.68, 2.75)				
2851	PENTANE	LL	0.52	(0.42, 0.68)	0.83	(0.57, 0.70)				
2973	PICOLINE, 2-	LL	0.43	(0.38, 0.50)	1.18	(1.08, 1.28)				
	PROPANE	G	0.95 I	(0.48, 0.72)	0.60 P	(0.59, 0.88)				
3063	PROPIONALDEHYDE	LL	1.14	(1.00, 1.32)	1.71	(1.11, 3.08)				
3068	PROPIONIC ACID	LL	1.30	(1.03, 1.70)	5.08 D	(0.73, >100.00)				
3070	PROPYL ALCOHOL	LL	0.93	(0.77, 1.18)	1.74	(1.08, 3.50)				
	PROPYLBENZENE, N-	LL	0.51	(0.45, 0.58)						
3090	PROPYLBENZENE	LL	0.77	(0.44, 2.66)	1.74 I	(0.15, 20.30)				
3120	PROPYLENE	G	0.83	(0.74, 0.95)	1.15	(0.89, 2.46)				
	PROPYLENE OXIDE	LL	0.47	(0.40, 0.55)	1.18	(1.03, 1.34)				
3130	PYRIDINE	LL	4.22	(3.45, 5.27)						
3230	STYRENE	LL	4.83 D	(1.24, >100.00)	8.91	(3.14, 22.50)				
3290	TETRACHLOROETHANE, 1,1,1,2	LL	7.89	(5.01, 13.80)	25.40	(8.08, >100.00)				
3291	TETRACHLOROETHANE, 1,1,2,2	LL	2.97	(1.71, 8.11)						
2890	TETRACHLOROETHYLENE	LL	0.39	(0.36, 0.43)						
3348	TOLUENE	HL	1.21 I	(0.50, 2.94)	2.68 D	(0.79, >100.00)				
3393	TRICHLOROBENZENE, 1,2,4-	LL	0.80	(0.72, 0.90)	0.47 I	(0.32, 0.68)				
3395	TRICHLOROETHANE, 1,1,1-	LL	1.25	(1.05, 1.50)	2.40	(1.81, 3.35)				
3400	TRICHLOROETHANE, 1,1,2-	LL	0.95	(0.83, 1.08)	3.69	(2.77, 5.18)				
3410	TRICHLOROETHYLENE	LL	0.95	(0.83, 1.08)	3.93	(2.68, 8.32)				
3420	TRICHLOROPROPANE, 1,2,3-	LL	0.98	(0.84, 1.78)	1.99	(1.27, 3.82)				
3450	TRIETHYLAMINE	LL	0.51	(0.40, 0.70)	1.48	(0.88, 2.78)				
3510	VINYL ACETATE	LL	1.27	(0.95, 1.82)	5.91 D	(1.26, >100.00)				
3520	VINYL CHLORIDE	G	0.84	(0.81, 1.38)	1.08	(0.59, 4.60)				

* ORGANIC CHEMICAL PRODUCERS DATA BASE

** G-GAS; LL-LIGHT LIQUID; HL-HEAVY LIQUID.

DEFINITION OF EXPLANATORY DATA CODES:

I INVERSE ESTIMATION METHOD
 D POSSIBLE OUTLIERS IN DATA
 N NARROW RANGE OF DATA
 X NO DATA AVAILABLE
 B 10,000 PPBV RESPONSE UNACHIEVABLE
 P SUSPECT POINTS ELIMINATED

TABLE 2-1 (Continued)

OCPDB* ID NO	COMPOUND NAME	VOLATILITY CLASS**	RESPONSE FACTORS WITH 95 % CONFIDENCE INTERVALS ESTIMATED AT 10,000 PPBV RESPONSE				TLV	
			OVA	CONFIDENCE INTERVALS	RESPONSE FACTOR	CONFIDENCE INTERVALS	RESPONSE FACTOR	CONFIDENCE INTERVALS
	VINYL PROPIONATE	LL	1.00 I (	0.57, 1.74)	1.21 I (	0.48, 3.20)		
3530	VINYLDENE CHLORIDE	LL	1.12 (	0.87, 1.52)	2.41 (	1.82, 3.35)		
3570	XYLENE, P-	LL	2.12 (	1.71, 2.68)	7.87 (	3.49, 24.90)		
3550	XYLENE, M-	LL	0.40 (	0.38, 0.48)	5.87 D (	0.81, >100.00)		
3580	XYLENE, O-	LL	0.43 (	0.28, 0.85)	1.40 (	0.61, 9.33)		

* ORGANIC CHEMICAL PRODUCERS DATA BASE

** G-GAS; LL-LIGHT LIQUID; HL-HEAVY LIQUID.

DEFINITION OF EXPLANATORY DATA CODES:

I	INVERSE ESTIMATION METHOD	X	NO DATA AVAILABLE
D	POSSIBLE OUTLIERS IN DATA	B	10,000 PPBV RESPONSE UNACHIEVABLE
N	NARROW RANGE OF DATA	P	SUSPECT POINTS ELIMINATED

Most of the response factors and associated confidence intervals were calculated using the classical method; those computed using the inverse method are noted in Table 2-1 with the explanatory code "I". Other explanatory codes used in Table 2-1 indicate data availability, data applicability and possible data uncertainties such as the presence of outliers.

In most cases the data fell very nearly on a straight line, but sometimes they did not. When outliers were clearly defined they were removed from the data and the response factor recomputed. This could not always be done. Often it was not clear which data points were valid and which were not. Additional data would be necessary to clarify these situations. Since it was not clear which points should be removed in these cases, all points were included.

Table 2-2 lists chemicals tested which do not appear to respond at a 10,000 ppmv reading at any concentration. The data and the fitted line showed that the response appears to be well below 10,000 ppmv for feasible actual concentrations. Sometimes it was not possible to distinguish between this situation and the problem of unresolved outliers mentioned above. Questionable or borderline cases are included in Table 2-1 rather than Table 2-2. Some of these chemicals have large estimated response factors and wide confidence intervals.

TABLE 2-2. A LIST OF TESTED CHEMICALS WHICH APPEAR TO BE UNABLE TO ACHIEVE AN INSTRUMENT RESPONSE OF 10,000 PPMV AT ANY FEASIBLE CONCENTRATION

OVA		TLV	
OCPDB*	Compound Name	OCPDB	Compound Name
-	Acetyl-1-propanol, 3-	120	Acetophenone
790	Carbon disulfide	-	Acetyl-1-propanol, 3-
810	Carbon tetrachloride	130	Acetylene
-	Dichloro-1-propanol,2,3-	160	Acrylic Acid
-	Dichloro-2-propanol,1,3-	360	Benzaldehyde
-	Diisopropyl benzene,1,3-	450	Benzonitrile
-	Dimethylstyrene,2,4	490	Benzoyl chloride
1221	Freon 12	530	Benzyl chloride
2073	Furfural	-	Butylbenzene, tert-
-	Methyl-2,4-Pentanediol,2-	810	Carbon tetrachloride
1660	Monoethanolamine	930	Chloroform
2770	Nitrobenzene	1040	Crotonaldehyde
2910	Phenol	1060	Cumene
-	Phenyl-2-Propanol,2-	1130	Cyclohexanol
		-	Dichloro-1-propanol,2,3-
		-	Dichloro-2-propanol,1,3-
		-	Diisopropyl benzene,1,3-
		-	Dimethylstyrene,2,4-
		2060	Formic acid
		1221	Freon 12
		2073	Furfural
		2200	Isobutylene
		-	Methyl-2,4-pentanediol
		2690	Methylstyrene, A-
		1660	Monoethanolamine
		2910	Phenol
		-	Phenyl-2-propanol,2-
		-	Propylbenzene,N-
		3230	Styrene
		2860	Tetrachloroethylene

*Organic Chemical Producers Data Base ID Number

SECTION 3

DATA ANALYSIS

The estimated response factors at 10,000 ppmv for the OVA and TLV were calculated as follows: A classical least squares line was fitted to the actual concentrations and instrument response readings on a logarithmic scale. The logarithm of the actual concentration at response log 10,000 was then computed by solving the equation using the fitted parameters. The result was then transformed to the data scale and divided by 10,000 to yield the desired response factor. A 95 percent confidence interval was computed following Brownlee (1965). The results are presented in Table 2-1.

In some cases Brownlee's (1965) confidence interval gives indeterminate values. This generally is due to a regression line that is flat relative to the observed error or to other data problems. Response factors estimated by an inverse regression were generally agreeable to the method described above. In those cases where the confidence interval was indeterminate, the inverse regression estimate and confidence interval were computed as an alternate. These inverse estimates are indicated in Table 2-1 by an explanatory code "I".

3.1 Classical Estimation of Response Factors

Let c_i be the actual concentration of the selected chemical and r_i be the corresponding instrument response for $i = 1, \dots, n$ observations. Consider then the transformed variates

$$y_i = \ln r_i$$

$$x_i = \ln c_i$$

and the relationship

$$y_i = \alpha + \beta x_i + e_i \quad (1)$$

where e_i is a random error of measurement. Let a and b be the classical least squares estimates of α and β , respectively.

Now, suppose that at some unknown x_0

$$\ln 10,000 = \alpha + \beta x_0$$

Then estimate x_0 by

$$\hat{x}_{01} = (\ln 10,000 - a)/b$$

and estimate the actual concentration corresponding to 10,000 ppmv observed instrument response by

$$\hat{c}_{01} = \exp(\hat{x}_{01}).$$

A confidence interval based on classical least squares estimation is derived by Brownlee (1965):

$$CL_k = \bar{x} + \frac{b(\phi - a)}{b^2 - t_k^2 s^2 / \sum (x_i - \bar{x})^2} - \frac{t_k s}{b^2 - t_k^2 s^2 / \sum (x_i - \bar{x})^2} \\ \times \left\{ \left[b^2 - \frac{t_k^2 s^2}{\sum (x_i - \bar{x})^2} \right] \left(\frac{1}{n} \right) + \frac{(\phi - a)^2}{\sum (x_i - \bar{x})^2} \right\}^{1/2}$$

where $k = 1, 2$

$t_1 = +t$ and $t_2 = -t$, with t being the tabular value for 95 percent confidence from student's t -distribution with $n-2$ degrees of freedom,

$s^2 = \sum(y_i - \hat{y}_i)/(n-2)$, with $\hat{y}_i = a + bx_i$, and

$\phi = \ln 10,000$ (In Brownlee, this is the mean of m observed responses. Here, the desired response is exact, so $m = \infty$ and $1/m = 0$).

The value CL_2 provides the lower limit and CL_1 provides the upper limit. These limits were computed in the logarithmic scale and then exponentiated into the data scale.

3.2 Inverse Estimation of Response Factors

Let x_i and y_i be as described above except that the inverse model relationship

$$x_i = \delta + \gamma y_i + \epsilon_i \quad (2)$$

is fitted by least squares. Let d and g be the estimates of δ and γ , respectively. Then if

$$x_o = \delta + \gamma \cdot \ln 10,000$$

estimate x_o by

$$\hat{x}_{o2} = d + g \cdot \ln 10,000.$$

Assuming normality of the errors ϵ_i ,

$$\hat{c}_{o_2} = \exp(\hat{x}_{o_2} + \frac{s^2}{2}) \quad (3)$$

is an approximately unbiased estimate of

$$c_o = \exp(x_o)$$

where s^2 is the estimated variance of the ϵ_i .

A confidence interval for x_o can be computed by the usual method:

$$\hat{x}_{o_2} \pm ts \sqrt{\frac{1}{n} + (x_o - \bar{x})^2 / \sum (x_i - \bar{x})^2}$$

where t is the 95 percent confidence level tabular value from Student's t -distribution with $n-2$ degrees of freedom. The confidence limits for $\hat{c}_{o_2}$ are computed by applying the transformation of (3) to these limits for x_o .

3.3 Discussion

Both of the models (1) and (2) cannot be correct simultaneously. If the model (2) is correct for the data, then the second or inverse approach is appropriate. However, if the model (1) is correct, then both approaches produce biased estimates. Generally speaking, the bias of the classical method will be smaller since it is a consistent estimator (bias reduces to zero as sample size increases without bound). The variance, however, may tend to be larger.

The model (1) appears to best approximate the data collected for this study. The classical approach was used as the primary method in this report

because of its consistency and reputed superiority in extrapolation. The properties of the data for most of the chemicals are such that the two approaches are virtually equivalent. A comparison of computed values confirmed this hypothesis.



SECTION 4

LITERATURE REVIEW

The statistical literature is clear on how to estimate y for a given x when y and x are related by

$$y = \alpha + \beta x$$

and y is subject to error. There has been much disagreement on how to estimate an unknown x given an observed value of y in the same situation. For many years it was universally assumed that the unknown x should be solved for using the usual parameter estimates. Some have suggested that the parameters be fitted by an inverse " x on y " regression and the unknown x estimated directly. There has been considerable controversy over this and some classic assumptions have proven unfounded.

The properties of the two estimates are summarized in the following table:

Property	Classic (y on x)	Inverse (x on y)
Bias	Biased	Biased
Asymptotic Bias	Consistent	Inconsistent
Variance		
--Under Normality	Infinite	Finite
--Truncated Normality	Large but Finite	Finite
Quality as Estimator		
--Well determined line	Both good and equivalent	
--Interpolation with ill determined line	Poor	Good
--Extrapolation	Good	Poor

In this study, many of the needed estimates are extrapolations just beyond the end of the available data. Many of the lines are fairly well determined. Thus, it is not completely clear which estimate is truly best here, but they are likely to be nearly equivalent. For this study, the classic estimator is used except in cases where it is clearly not appropriate. The inverse estimator is used for these cases.

The following is a summary of the literature examined on this subject:

The relationship between two variables y and x can often be expressed as a simple linear function. If the variable x takes on fixed values x_i for which responses y_i are measured, the relationship is usually expressed as the linear regression model

$$y_i = \alpha + \beta x_i + e_i \quad i = 1, \dots, n \quad (1)$$

where the e_i are random errors and n is the number of observed pairs (x_i, y_i) . It is usually assumed that the e_i are independent and identically distributed, and furthermore, that they are distributed according to a normal distribution with mean 0 and variance σ^2 .

It is well known that the usual least squares estimates of α and β in (1) are Best Linear Unbiased Estimators (BLUE), where "Best" means "smallest variance". Estimates of $y_i = \alpha + \beta x_i$ based on the least squares estimates of α and β are also BLUE. Under the assumption of normality of errors, these estimates are also maximum likelihood estimates. The least squares estimators have many other useful and beneficial properties.

Here we must consider a slightly different problem known as the Calibration Problem. In addition to the responses y_i on known x_i , there are further responses z_j corresponding to an unknown x_0 . The extended

model is

$$y_i = \alpha + \beta x_i + e_i \quad i = 1, \dots, n \quad (2)$$

$$z_j = \alpha + \beta x_0 + e_j \quad j = 1, \dots, m \quad (3)$$

The problem here is to provide point and interval estimates for x_0 rather than to estimate y_i, α , or β .

The above discussion closely follows Kalotay (1971).

At least since Eisenhart (1939) the accepted solution has been to estimate the unknown x_0 by inverting the least squares estimates from (2):

$$\hat{x}_{01} = \frac{\bar{z} - a}{b}$$

where $\bar{z}$ is the mean of the z_j in (3) and a and b are the least squares estimates of α and β respectively. This is referred to as the "classical" estimator.

Krutchkoff (1967, 1969) challenged this approach by claiming that a superior estimate could be had by inverting the regression variables; that is, compute the least squares estimates of δ and γ from the model

$$x_i = \delta + \gamma y_i + \varepsilon_i$$

and estimating the unknown x_0 directly by

$$\hat{x}_{02} = d + g \bar{z}$$

where d and g are the least squares estimates of δ and γ respectively.

This is referred to as the "inverse" estimator. Krutchkoff based his assertion not on mathematical statistical properties but on Monte Carlo simulations.

After Krutchkoff's results appeared there was a flurry of reaction. Williams (1969) reported that if normality of errors was assumed, then the classical estimator $\hat{x}_{o_1}$ has infinite variance. He showed further that no unbiased estimate of x_o would have finite variance under this condition and that the inverse estimator $\hat{x}_{o_2}$ is biased. On this basis he recommends the classical estimator.

Berkson (1969) showed that the inverse estimator $\hat{x}_{o_2}$ has a smaller Mean Square Error (MSE) than the classical estimator $\hat{x}_{o_1}$ when x_o is in the neighborhood of $\bar{x}$, the mean of the x_i 's. However, if x_o is not near $\bar{x}$ then the classical estimator has lower MSE. Furthermore, the inverse estimator is not only biased for finite n but is inconsistent: The bias does not reduce to zero as n increases without bound. On this basis he also recommends the classical estimator. Martinelle (1970), looking at relative efficiency, agrees with Berkson. Saw (1970) agrees with Martinelle on the same basis.

Hoadly (1970) reaffirms that the classical estimator has infinite variance if normality is assumed. It may also have infinite confidence intervals. He then proposes an alternative solution based on Bayesian methods.

Williams (1969) had suggested that an MSE comparison was inappropriate since under normality the classical estimator always has infinite MSE so any finite MSE estimate could beat it, including any constant value. To avoid this Halperin (1970) compares the two approaches based on Pitman's closeness criterion. He showed that the inverse estimator was superior in the neighborhood of $\bar{x}$. However, he believes that the neighborhood where the inverse estimator is superior is too small to recommend it over the

classical estimator.

Krutchkoff (1970, 1971, 1972) defended his point of view primarily with Monte Carlo simulations. His results indicate that the inverse estimator is better by the closeness criterion when x_0 is near $\bar{x}$. The classical estimator is better further out from $\bar{x}$.

Kalotay (1971), recognizing that the classical estimator has certain unsatisfactory properties, proposes an alternative estimator which he compares to Hoadley's (1970) Bayesian estimator. Kalotay's structural solution to the problem is not practical for any but very small sample sizes. He points out that under certain conditions his estimator reduces to the inverse estimator.

Shukla (1972) developed approximate formulas for bias, variance, and MSE of both classical and inverse estimators. He shows that both are biased but that the classical estimator is asymptotically unbiased, unlike the inverse estimator. However, both biases reduce to zero when the unknown x_0 is at the mean of the x_i . Using the optimal design for the classical estimator, the inverse estimator has smaller MSE than the classical estimator in the range of the x_i . Nevertheless, Shukla (1972) recommends the classical estimator for its bias properties.

Earlier critics of the inverse estimator rejected it because it is not only biased but inconsistent. The classical estimates for α , β , and y_i are unbiased but it was not proven that the classical estimator $\hat{x}_{01}$ is unbiased for x_0 . Naszodi (1978) shows that the classical estimator is also biased. Furthermore, if normality is assumed, it has no expectation.

Naszodi (1978) points out that in practical problems, perfect normality rarely exists. In fact, even when data errors are normal, extreme values will generally be discarded. Thus, in reality the distribution is a truncated normal. In any case, Naszodi (1978) gives a new estimator

involving an approximate bias correction for the classical estimator. Based on a simulation, all three estimators appear to have about the same MSE with the inverse estimator having the lowest and the classical the highest.

Brown (1979) points out that the objectives of the estimation in (1) and in the calibration problem (2) and (3) are different. There is no argument that the classical approach is "best" for estimating α and β . In the calibration problem, the objective is the estimation of the unknown x_0 .

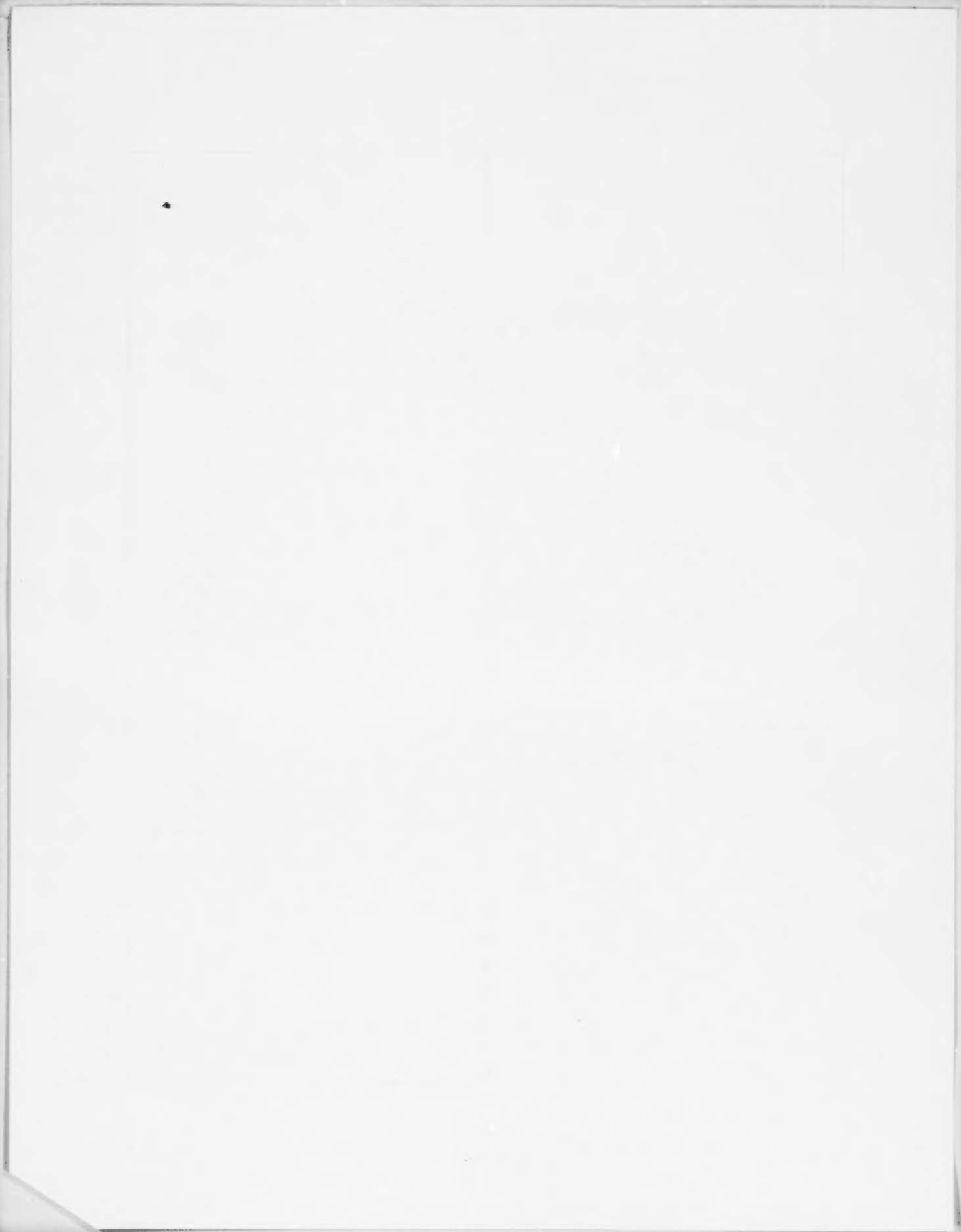
Brown's (1979) initial approach is to find the minimum MSE linear estimator of x_0 . He shows, however, that the MSE of a linear estimator depends on x_0 . To cope with this he proposes an unspecified distribution of future x_0 's to be estimated. After integrating the MSE over this distribution, a new minimum integrated MSE estimator (IMSE) is obtained.

Brown (1979) does not propose his minimum IMSE estimator for general use, although he notes that it could be used in some cases. He uses it as a benchmark to compare the classical and inverse estimators. When the regression line is well determined by the data, the two estimators are virtually equivalent. When the regression line is ill determined, the classical estimator is very bad. The inverse estimator will be very good if the mean of the calibration sample $\bar{x}$ is near the expected value of the future unknown x_0 's. If the range of future x_0 's is large relative to the range of the calibration x_1 's, then the inverse estimator is very bad while the classical estimator is very good.

The previous discussion applies to point estimates. Much less is said about interval estimates for confidence intervals. Williams (1969) suggests that the two point estimators should be compared on the basis of confidence intervals. Shulka (1972), referring to this, states that it is not clear how a confidence interval can be constructed for the classical estimate. The usual method of estimating a confidence interval for a

predicted response may be applied in reverse to give a confidence interval for the inverse estimator.

Brownlee (1965) gives a confidence interval based on the classical estimation method. Naszodi (1978) refers to this interval without comment. Hoadley (1970) comments that a confidence interval for the classical estimator may be infinite. Trout and Swallow (1971) describe Scheffe confidence bounds for individual observations. They then derive uniform confidence bands over a specified range of x .



REFERENCES

- Berkson, J. (1969). Estimation of a Linear Function for a Calibration Line: Consideration of a Recent Proposal. Technometrics, 21, 649.
- Brown, G. E., D. A. DuBose, W. R. Phillips, and G. E. Harris, "Response Factors of VOC Analyzers Calibrated with Methane for Selected Organic Chemicals", EPA Report No. 600/2-81-002 (NTIS NO. PB81-136194), Research Triangle Park, N.C., 1980.
- Brown, G. H. (1979). An Optimization Criterion for Linear Inverse Estimation. Technometrics, 21, 575.
- Brownlee, K. A. (1965). Statistical Theory and Methodology in Science and Engineering, New York: John Wiley & Sons, Inc.
- Eisenhart, C. (1939). The Interpretation of Certain Regression Methods and Their Use in Biological and Industrial Research. Annals of Mathematical Statistics, 10, 162.
- Environmental Protection Agency, VOC Fugitive Emissions in Synthetic Organic Chemicals Manufacturing Industry--Background Information for Proposed Standards, Emission Standards and Engineering Division, Research Triangle Park, N.C., March, 1980 (Draft).
- Halperin, M. (1970). On Inverse Estimation in Linear Regression. Technometrics, 12, 727.
- Hoadley, B. (1970). A Bayesian look at Inverse Linear Regression. J. Amer. Statist. Assoc., 65, 356.
- Kalotay, A. J. (1971). Structural Solution to the Linear Calibration Problem. Technometrics, 13, 761.
- Krutchkoff, R. G. (1967). Classical and Inverse Regression Methods of Calibration. Technometrics, 9, 425.
- Krutchkoff, R. G. (1968). Letter to Editor. Technometrics, 10, 430.
- Krutchkoff, R. G. (1969). Classical and Inverse Regression Methods of Calibration in Extrapolation. Technometrics, 11, 605.
- Krutchkoff, R. G. (1970). Letter to Editor. Technometrics, 12, 433.
- Krutchkoff, R. G. (1971). The Calibration Problem and Closeness. Journal of Statistical Computation and Simulation I, 87.
- Krutchkoff, R. G. (1972). Letter to Editor. Technometrics, 14, 241.

- Martinelle, S. (1970). On the Choice of Regression of Linear Calibration. Comments on a paper by R. G. Krutchkoff. Technometrics, 12, 157.
- Naszodi, L. J. (1978). Elimination of the Bias in the Course of Calibration. Technometrics, 20, 201.
- Saw, J. G. (1970). Letter to Editor. Technometrics, 12, 937.
- Shuckla, G. K. (1972). On the Problem of Calibration. Technometrics, 14, 547.
- Trout, J. R. and W. H. Swallow (1979). Regular and Inverse Interval Estimation of Individual Observations Using Uniform Confidence Bands. Technometrics, 21, 567.
- Williams, E. J. (1969). A Note on Regression Methods in Calibration. Technometrics, 11, 189.